

**UNITED STATES DISTRICT COURT
SOUTHERN DISTRICT OF NEW YORK**

WENDY GRAHAM, MARIA HÖHN, MIA
MASK, CINDY SCHWARZ, and DEBRA
ZEIFMAN on behalf of themselves and all others
similarly situated,

Plaintiffs,

v.

VASSAR COLLEGE,

Defendant.

No. 7:23-cv-07692

**STIPULATED ORDER REGARDING
THE PRODUCTION OF
ELECTRONICALLY STORED
INFORMATION (“ESI”) AND HARD
COPY DOCUMENTS**

I. PURPOSE

This Stipulation and Order (“Order”) regarding the production of Electronically Stored Information (“ESI”) and Hard Copy Documents (collectively, “Document” or “Documents”) shall govern discovery of Documents in this case as a supplement to the Federal Rules of Civil Procedure, and any other applicable orders and rules. Nothing in this Order is intended to expand or limit the Parties’ obligations under the Federal Rules of Civil Procedure. Any disputes arising out of the production of Documents subject to this Order shall be resolved according to Federal Rules of Civil Procedure, the Local Rules, and the Court’s Standing Orders.

II. COOPERATION

The Parties are aware of the importance the Court places on cooperation and commit to cooperate in good faith throughout the matter consistent with this Court’s Guidelines for the Discovery of ESI.

III. LIAISON

The Parties have identified liaisons to each other who are and will be knowledgeable about and responsible for discussing their respective ESI. Each e-discovery liaison will be, or have access to those who are, knowledgeable about the technical aspects of e-discovery, including the location, nature, accessibility, format, collection, search methodologies, and production of ESI in this matter. The Parties will rely on the liaisons, as needed, to confer about ESI and to help resolve disputes without court intervention.

IV. PRESERVATION AND IDENTIFICATION OF ESI

A. The Parties will discuss their preservation obligations and needs and agree that preservation of potentially-relevant ESI will be reasonable and proportionate. The Parties will exchange a list of the types of ESI they believe should be preserved and the custodians, or general job titles or descriptions of custodians, for whom they believe ESI should be preserved, e.g., “HR head,” “scientist,” and “marketing manager.” The Parties shall add or remove custodians as reasonably necessary. The Parties will seek agreement on the identity of custodians for whom ESI will be preserved and will raise any disputes regarding these custodians for resolution by the Court.

B. To reduce the costs and burdens of preservation and to ensure proper ESI is preserved, the Parties agree that:

1. Within 30 days following receipt of the Requesting Party’s requests for production of Documents, or a later time if agreed to by the Parties, the receiving Party shall disclose the names of individuals and/or sources from which the Producing Party plans to collect Documents in response to the document requests. The custodians shall be identified by name,

title, and connection to the instant litigation.

2. Where applicable, sources of Documents include but are not limited to desktop servers, computers, laptops, local hard drives, file servers, databases, SANs, NASs, email servers, web servers, network home or personal file shares, any shared drives, document management systems (DMS), record management systems (RMS), content management systems (CMS), departmental/project/collaborative/shared storage spaces removable storage, email, removable storage media, on-line data storage (e.g., Dropbox or Google Drive), messaging phones, tablets social media, and physical or hard copy files and Documents.

3. The Parties shall have a continuing obligation to take reasonable and proportional steps to identify and preserve custodial and non-custodial data sources that may contain information that is relevant to the claims and defenses in this litigation.

4. If a Producing Party is aware of not reasonably accessible data that is likely to contain unique, discoverable ESI, it will identify the source from which it claims the data is not reasonably accessible to the Receiving Party.

5. Examples of data the Parties agree not to preserve are:

- a. Deleted or unallocated data only accessible by forensics;
- b. Random access memory (RAM), temporary files, or other ephemeral data that are difficult to preserve without disabling the operating system;
- c. Structural files not material to individual file contents (e.g. .CSS, .XSL, .XML, .DTD, etc.);
- d. Server, system, or network logs;
- e. Back up tapes or other storage media used for disaster recovery

purposes; and

- f. Data in metadata fields that are unrelated to user interaction and instead updated only automatically (e.g., in connection with a virus scan or software update).

6. To the extent the Producing Party identifies other forms of potentially responsive ESI whose preservation requires extraordinary affirmative measures that are not utilized in the ordinary course of business, the Producing Party shall promptly request to meet and confer with the Receiving Party regarding whether that form of ESI should be preserved, and will preserve that form of ESI until that meet and confer is complete.

V. SEARCH & REVIEW

1. The Parties will meet and confer to discuss the use of search terms, filters and date ranges or the use of advanced search and retrieval technologies. A Producing Party shall specify the technique or techniques it will use to search for material it reasonably anticipates will be responsive to Requests for Production. If the Requesting Party objects to the particular technique or techniques, the Parties will timely meet and confer regarding the matter and submit any unresolved disputes to the Court for resolution.

2. A Party who intends to use search terms will describe a Search Protocol that will include: (a) the criteria to be used to identify the universe of Documents to which search terms will be applied (e.g., date range), and (b) proposed search terms to be applied to that universe of Documents, subject to revision based on, for example, meet and confer of the Parties and the number of hits returned by each term. A Requesting Party may also suggest search terms to be applied. The Parties will work in good faith to negotiate a set of search terms that the Producing Party will utilize to identify documents for review (“Search Term Review Population”), subject

to meet and confer negotiations with the Requesting Party and Sections V(4) – V(6) below (“Agreed Search Terms”).

3. To the extent a Producing Party identifies documents for review outside of the Agreed Search Terms, the Producing Party waives any right to later claim any burden associated with search efforts using those terms.

4. Any Search Protocol shall include a requirement for the Producing Party to review one random sample of 2,000 distinct Documents that do not hit on the Agreed Search Terms (the “Null Set”), and to produce to the Requesting Party any responsive, non-privileged Documents identified. The Null Set of 2,000 distinct Documents shall be generated after all document processing is substantially complete. These 2,000 Documents should come from each custodian and include Documents throughout the entire date range. Producing Party should then review the 2,000 Documents that do not hit on any terms and produce any responsive Documents. If more than 10% of these Documents are responsive (i.e., 200 Documents), then Parties shall revisit the agreed-upon search terms by analyzing the Documents produced and determining which additional terms are necessary. The Parties agree to meet and confer regarding the Search Protocol and to raise any disputes regarding the Search Protocol for resolution by the Court.

5. A Producing Party may use technology assisted review¹ (“TAR”) to prioritize the review of their Search Term Review Population. However, if the Producing Party reasonably believes it necessary to exclude a subset of the Search Term Review Population from manual review to avoid undue burden and cost, the Producing Party will notify the Receiving Party and

¹ For purposes of this Protocol, TAR is limited to predictive coding or continuous active learning technologies and does not include de-duplication or email threading.

the Parties agree to meet and confer regarding the potential exclusion of documents from manual review and to raise any disputes regarding that process for resolution by the Court.

6. A Party who intends to use TAR for purposes other than to prioritize the review of their Search Term Review Population will describe a TAR Protocol that will include: (a) the TAR system to be used; (b) the criteria to be used to identify the universe of Documents to which TAR is to be applied (the “TAR Universe”); (c) its methodology for training the TAR model and identifying responsive Documents; (d) the methodology for validation testing to be used; (e) the subject-matter expert who will oversee the implementation of the TAR Protocol; and (f) the process by which Documents excluded as not conducive to categorization (e.g. multimedia files, primarily numerical spreadsheets, database files) will be reviewed for production.

7. Documents which are reasonably believed to be responsive to discovery requests and for which text-based search technologies are fundamentally ineffective must be reviewed without culling by search terms, TAR, or other technologies that rely primarily on text.

VI. PRODUCTION FORMATS

The Parties agree to produce Documents in the formats set forth in Appendix 1 and 2 hereto.

VII. MODIFICATION

This Order may be modified by a stipulation of the Parties or by the Court for good cause shown.

VIII. MISCELLANEOUS PROVISIONS

A. Objections Preserved.

Nothing in this Order shall be interpreted to require disclosure of relevant information protected by the attorney-client privilege, attorney work-product doctrine, or any other

applicable privilege or immunity. Except as provided expressly herein, the Parties do not waive any objections as to the production, discoverability, authenticity, admissibility, or confidentiality of Documents and ESI.

B. Technical Variances.

Recognizing that each Producing Party may experience production issues due to data systems or files that may not be fully compatible with the technical specifications set forth herein, any practice or procedure set forth herein may be varied by agreement of the Parties, confirmed in writing, to accommodate such issues and/or where such variance is deemed appropriate to facilitate the timely and economical production of Documents or ESI. No Party shall unreasonably object to any such variance.

C. Privileged Documents and Privilege Log.

To the extent that a Party reasonably determines that one or more responsive Documents are not discoverable or a portion of a Document needs to be redacted because it is subject to the attorney-client communication privilege or work product doctrine, or otherwise not discoverable on the basis of a recognized protection or privilege (collectively, the “Privileges” and each a “Privilege”), the Party shall, within 30 days of the production from which Documents are withheld or redacted for privilege, produce a log treating each Document withheld or redacted for privilege from that production separately that sets forth: (a) the nature of the privilege(s) or other protection claimed; (b) the date of the Document; (c) the name of the author(s)/addresser, addressee(s) and all recipients of the Document and whether any person identified is an attorney or an employee of any Defendants’ legal department; (d) a description of the general subject matter contained in the Document and the type of Document (e.g., letter, memorandum, handwritten notes) sufficient to allow the Receiving Party to assess the claimed Privilege and/or

to allow the Court to rule upon the applicability of the claimed protection; and (e) the custodian and/or location of the Document. For redacted Documents with multiple redactions, one log entry can be used to describe the multiple redactions, so long as a single entry is sufficient to adequately identify the basis for all redactions, as described in this paragraph. Names of counsel shall be designated with an asterisk.

The following need not be placed on a privilege log: (1) communications regarding this matter that are exclusively between Outside Counsel for Defendant and Vassar College employees Shay Humphrey, Elizabeth Bradley, and/or their Administrative Assistants, and (2) communications regarding this matter that are exclusively between Outside Counsel for Plaintiffs and any member of the proposed Class who sought their legal advice in this matter. Outside Counsel for Defendant includes attorneys or staff at Seyfarth Shaw LLP. Outside Counsel for Plaintiffs includes attorneys or staff at Lieff Cabraser Heimann & Bernstein, LLP or Equal Rights Advocates.

D. Hard Copy Document Storage.

During the pendency of this litigation, the Parties shall make reasonable efforts to preserve the originals of all hard copy Documents as to which there may be issues of legibility of all or any part of the production copy. Each Party reserves the right to request to inspect such original Documents of the opposing Party or Parties, which request shall not be unreasonably denied. If such request to inspect is denied, the Party may seek relief from the Court.

E. Short Message Data.

Electronic messages exchanged between users on communication software such as Microsoft Teams and Slack shall be produced in a searchable format that preserves the conversational relationship and presentational features of the original messages, such as emojis,

images, video files, and animations. Electronic messages must not be converted to unitized files that contain less than a 24-hour period of conversation. Redactions may be applied to privileged or non-responsive portions of a conversation. To the extent electronic messages cannot be produced in a reasonably useable format, the Parties will meet and confer to address the identification, production, and production format of short message data.

F. **Pointer Hyperlinks.**

The Parties agree that documents containing internal hyperlinks pointing to Documents on a system within a Producing Party's custody, possession, or control (e.g., such as an internal document management system, Google Drive, Google Docs, SharePoint, Office365, or similar document hosting and collaboration service) do not need to be produced in the first instance as part of the same family group as the Document residing at the location at which that Hyperlink points. If there are particular responsive Documents containing hyperlinks where the Document to which that hyperlink points cannot be located in the production, the Receiving Party may submit a list of the requested hyperlinks (identifying the underlying Documents by Bates number) to the Producing Party, and the Producing Party will engage in reasonable efforts to locate, collect, and review the Document at that pointed location and either identify it by Bates number or produce it consistent with this Protocol if it not already produced, responsive, and not privileged. After reviewing the listing of hyperlinked Documents provided by the Receiving Party, the Producing Party will inform the Receiving Party if it reasonably believes that locating, collecting, and reviewing the requested hyperlinked Documents would be unduly burdensome and the Parties will meet and confer, if necessary.

G. **Collaboration Tools.**

The parties agree to produce the latest in time version of documents stored on a collaboration or document management tool. To the extent the receiving party desires to receive previous versions of documents, if any, the receiving party will identify the document(s) with specificity and the parties will engage in a meet and confer process relating to the same.

H. **Reproduction as Natives.**

The Parties agree that to the extent any Party seeks production in native format of specifically identified ESI produced originally in TIFF form, the Producing Party shall respond reasonably and in good faith to any such request.

IT IS SO STIPULATED, THROUGH COUNSEL OF RECORD.

DATED: 01/28/2025 /s/ Michelle A. Lamy
Attorneys for Plaintiffs

DATED: 01/28/2025 /s/ Maria Papasevastos
Attorneys for Defendant

PURSUANT TO STIPULATION, IT IS SO ORDERED.

DATED: January 29, 2025 
The Honorable Victoria Reznik
United States Magistrate Judge

APPENDIX 1
PRODUCTION DELIVERY REQUIREMENTS

I. GENERAL PRODUCTION PROVISIONS

The Parties have agreed that ESI should be produced as TIFF images and in Native Format where applicable with accompanying data and image load files.

A. TIFF Image Files.

The Parties agree that all Documents, with the exception of Documents produced in Native Format, will be produced as single-page black and white Group IV TIFF image files of at least 300 dpi resolution with 1 bit depth. Page size shall be 8.5 x 11 inches, unless in the reasonable judgment of the Producing Party, a particular item requires a different page size. Each image file will use the Bates number of the page as its unique file name. Original document orientation as displayed in the native file should be maintained in the TIFF image (i.e., portrait to portrait and landscape to landscape).

B. Text Files.

Each Document produced under this Order shall be accompanied by a Document level text file containing all of the text for that Document, not one text file per page. Each text file shall be named to use the Bates number of the first page of the corresponding Document. The text files shall be generated by extracted text from native files and OCR text files from hard copy scanned Documents as applicable. To ensure optimal accuracy and quality, OCR software used for extracting text from native files and for processing hard copy scanned Documents shall be set to the highest quality setting. The .DAT load file shall include a link to the corresponding text file.

C. OCR Text File.

The Parties will provide searchable OCR text of any paper or imaged Documents.

D. **Extracted Text Files from ESI.**

The Parties shall extract the text of each ESI item directly from the ESI native file, where extracted text is reasonably available. For contacts and calendars collected and/or processed after the execution date of this Order, fields should be extracted and produced as text.

E. **OCR Text for Redacted Documents.**

The Parties will provide searchable OCR text for any redacted files.

F. **Bates Numbering.**

1. Each TIFF image produced under this Order should be assigned a Bates number that must: (1) be unique across the entire document production; (2) maintain a constant length of nine numeric digits (including 0-padding) across the entire production; (3) contain only alphanumeric characters, no special characters or embedded spaces; and (4) be sequential within a given Document. The Producing Party will identify the Bates number range of each production in a cover letter or production log accompanying the production. If a Producing Party skips a Bates number or set of Bates numbers in a production, the Producing Party will identify and note the gap in the cover letter or production log accompanying the production.

2. The Producing Party will brand all TIFF images in the lower right-hand corner with its corresponding Bates number, using a consistent font type and size. If the Receiving Party believes that a Bates number obscures the content of a Document, then the Receiving Party may request that the Document be produced with the Bates number in a different position.

G. **Parent-Child Relationships.**

Parent-child relationships for all embedded ESI Documents (e.g., the association between an attachment and its parent email, or a spreadsheet embedded within a word processing

Document), must be preserved by assigning sequential Bates numbers to all items within the parent-child group, and identifying those Bates numbers in the relevant ESI metadata and coding fields specified in Appendix 2. For example, if a Party is producing an email with embedded attachments, the attachments must be processed and assigned Bates numbers in sequential order, following consecutively behind the parent email. Embedded images that are included in-line within the parent document need not be extracted and produced separately. All graphic embedded objects or embedded objects without content (company logos, social media icons, blank documents, etc.) may be excluded from production by the Producing Party provided they are either (1) first manually reviewed and confirmed to be graphic embedded objects or embedded objects without content, or (2) have a hash value identical to a graphic embedded object or embedded object without content first manually reviewed and confirmed to be such an embedded object, but will be produced upon reasonable request by the receiving party. A producing party is not required to review or produce extracted embedded OLE documents if the document they were extracted from will be produced in native format.

H. **Color Documents.**

If an original ESI Document contains color text, markings or graphics, and the Receiving Party believes it is necessary to view such Document in its original color to understand its full meaning or content, then the Receiving Party may request that the ESI Document be produced in color format. The Producing Party shall then reproduce such Document(s) and/or ESI in color JPEG format, or in Native Format. This section also applies to Documents that are produced as TIFF images. Color images should be produced as single page JPG files at 300dpi with JPG compression and a high-quality setting as to not degrade the original image. Redacted versions of

documents that would have been produced in color in their unredacted form shall be produced in color as detailed herein.

I. Confidentiality Designations.

If a particular Document has a confidentiality designation, the designation shall be stamped on the face of all TIFF images pertaining to such Document, in the lower left-hand corner of the Document, or as close thereto as possible while preserving the underlying image. If the Receiving Party believes that a confidentiality designation obscures the content of a Document, then the Receiving Party may request that the Document be produced with the confidentiality designation in a different position. No Party may attach to any filing or any correspondence addressed to the Court (including any Magistrate Judge), or any adverse or third Party, or submit as an exhibit at a deposition or any other judicial proceeding, a copy (whether electronic or otherwise) of any native format Document produced by any Party without ensuring that the corresponding Bates number and confidentiality legend, as designated by the Producing Party, appears on the Document. For each Document that is marked confidential or highly confidential – attorneys’ eyes only, a Confidentiality field will be populated with the word “Confidential” or “Highly Confidential – Attorneys’ Eyes Only” in the .dat file. Also, any Documents marked Confidential or Highly Confidential – Attorneys’ Eyes Only must be handled in accordance with the Protective Order entered in this case.

II. PRODUCTION OF HARD COPY DOCUMENTS

All hard copy Documents that are scanned will be produced in electronic form. Where necessary and practicable, hard copy Documents in color will be scanned in color to ensure full information is communicated in the scanned copy. Scanned color Documents will be provided in JPG file format.

The parties will not be required to rescan or re-unitize documents that were scanned or unitized prior to the entrance of this Order.

A. **Unitization of Paper Documents.**

To the extent practicable, hard copy Documents shall be unitized using logical document determinations or “LDD.”

B. **Identification.**

Where a Document or group of Documents has an identification spine, “post-it note,” or any other label, the information on the label shall be scanned and produced to the extent practicable.

C. **Custodian Identification.**

The Parties will utilize reasonable best efforts to ensure that paper records for a particular custodian or department level custodian, which are included in a single production, are produced in consecutive Bates stamp order.

D. **Metadata.**

The metadata associated with each hard copy Document need only identify the Bates number, the custodian associated with that hard copy Document, and any Confidential Designation or Redaction applied to that Document.

III. **PRODUCTION OF “ESI”**

A. **De-NISTING and System Files:**

ESI productions shall be de-NISTed using the industry standard list of such files maintained in the National Software Reference Library by the National Institute of Standards & Technology. De-NISTED files need not be produced. The Parties may mutually agree upon any additional file types that can be excluded from review and production. A Producing Party shall

identify any additional standard, readable, and reviewable file types which have been excluded from its document review population for any production made following the date of this Order, and will utilize reasonable best efforts to do the same with respect to any productions made prior to the date of this Order.

If a Party excludes from review a standard, readable, and reviewable file type not within the industry standard, that Party must disclose such exclusion to the other Parties.

B. Native Files:

Certain files types, such as presentation-application files (e.g., MS PowerPoint), spreadsheet-application files (e.g., MS Excel, .csv), and multimedia audio/visual files such as voice and video recordings (e.g., .wav, .mpeg, and .avi), shall be produced in Native Format. For files produced in Native Format, the Producing Party shall provide a single-page TIFF slip-sheet with the applicable Bates stamp and confidentiality designation indicating that a native item was produced. The corresponding load (.DAT) file shall include a NativeFileLink which provides the relative path linking information for each native file that is produced. In addition, the confidentiality designation will be indicated in the name of the native file where reasonably feasible.

C. Metadata Fields and Processing.

1. ESI shall be processed in a manner that preserves the source native file and relevant metadata without modification, including their existing time, date, and time-zone metadata consistent with the requirements provided in this Order. The Producing Party will provide information on the time zone the data was processed in to the Receiving Party.

2. Hidden text. ESI shall be processed, to the extent practicable, in a manner that preserves hidden columns or rows, hidden text, notes, or worksheets, speaker notes, tracked

changes, redlines and comments. Upon reasonable request, a Producing Party will produce files with any such information in Native Format.

3. Compressed Files and Encrypted Files. Compressed file types (i.e., .CAB, .GZ, .TAR .Z, .ZIP) shall be decompressed in a reiterative manner to ensure that a zip within a zip is decompressed into the lowest possible compression resulting in individual files. The Producing Party will take reasonable steps, prior to production, to unencrypt any discoverable electronically stored information that exists in encrypted format (e.g., because password-protected) and that can be reasonably unencrypted.

4. Metadata and Coded Fields. ESI items shall be produced with all of the metadata and coding fields set forth in Appendix 2.

This Order does not create any obligation to create or manually code fields that are not automatically generated by the processing of the ESI, or that do not exist as part of the original metadata of the Document; provided, however, that the Producing Party must populate, where possible, the (a) BegBates, (b) EndBates, (c) BegAttach, (d) EndAttach, (e) Custodian, (f) NativeFileLink fields, if applicable, (g) TextPath, (h) Redaction status, and (i) Confidentiality designation. These fields should be populated for all produced ESI, as well as paper Documents converted to electronic form, regardless of whether the fields can be populated pursuant to an automated process.

IV. DATABASES, STRUCTURED, AGGREGATED OR APPLICATION DATA

1. The Parties will produce Documents collected from databases or other structured databases in a reasonably useable format. If the data cannot be produced in a reasonably useable format, the Parties will meet and confer to address the production format. Prior to or during any such meet and confer, the Producing Party will provide sufficient information to enable the

Receiving Party to evaluate the best method and format of production. If the Parties cannot reach agreement, the matter will be decided by the Court or its designee.

V. RESPONSIVENESS, PRIVILEGE & REDACTIONS.

A. Responsiveness:

The Parties agree that even if an attachment to a responsive e-mail is wholly nonresponsive to the discovery requests, it will be produced along with any other Documents in the family.

B. Privilege:

To the extent that a Defendant reasonably determines that e-mail or related attachments that are responsive to Plaintiff's document request are not discoverable because they are subject to a Privilege, Defendant shall produce a log treating each e-mail and attachment withheld separately, subject to the exceptions in paragraph VIII (C) above. The emails within a thread of emails shall each be separately logged for privilege if withheld or redacted.

C. Redactions:

The parties agree that Excel files may be redacted in their native form. For non-Excel ESI items requiring redaction, they shall be produced in TIFF format with each redaction clearly indicated. In cases where TIFF format is not practicable, such items should be produced in redacted native format, as outlined below. Any unaffected data fields shall be provided. The redaction of any material for privilege or other reason shall be governed by the applicable rules and case law regarding privilege and the provisions contained in the Protective Order entered in this action. Relevance redactions are not permitted absent a meet and confer and agreement by both Parties, with the exception that redaction of personal information as defined by Federal Rule of Civil Procedure 5.2(a) is permitted for individuals who are not Plaintiffs in this action. If

the Receiving Party objects to a particular redaction, the Parties will timely meet and confer regarding the matter and submit any unresolved disputes to the Court for resolution. Documents that are to be produced in native format, but that require redactions may be produced as TIFF images with the relevant portions redacted, or if a TIFF image production is not practicable (e.g., the file is a video or very large spreadsheet), the producing party may produce a copy of the native file with the relevant portions replaced with “REDACTED” or a similar mark. If modification of a native file is required for redaction purposes, metadata information associated with that file should remain unchanged, unless it also requires redaction. For each document that is redacted, a Redaction field will be populated with the word “REDACTED” in the .DAT file. Also, the producing party will keep a pristine original copy of the native document.

D. Printing Specifications for Excel and PowerPoint files:

For Excel and PowerPoint type files that are printed to TIFF for redaction and redacted, the following printing options shall be enabled:

<p style="text-align: center;">Excel Print to TIFF Options</p> <ul style="list-style-type: none"> • Unhide columns and rows • Unhide worksheets • Autofit columns and rows, settings to be over by columns first and, then down by rows • Print gridlines • Do not apply Autofilter • Display headings • Display comments • Header and Footer filename field handling: Show field code
<p style="text-align: center;">PowerPoint Print to TIFF Options</p> <ul style="list-style-type: none"> • Print notes pages • Print hidden slides • Print comments

1. The Producing Party shall also make reasonable efforts to ensure that any spreadsheets produced only as TIFF images because of redactions are formatted so as to be

legible. For redacted items which were originally ESI, all metadata fields noted in this protocol that do not contain privileged or protected information will be provided and will include all non-redacted data. A Document's status as redacted does not relieve the Producing Party from providing all of the metadata required herein.

E. **De-duplication and Document Families:**

ESI may be de-duplicated globally at the parent level using MD5 hash or SHA-1 hash. For emails, de-duplication may be performed using a hash calculated using combination of metadata fields. Such hash shall, at a minimum, account for all sender and recipient metadata including BCC recipients, the email subject, date sent, and attachment names.

The Parties are permitted to produce the longest unique email chain and do not need to separately produce the lesser-included emails unless those lesser-included emails have bcc recipients or unique attachments not included in the longest chain. If a lesser-included email has a bcc recipient or unique attachment, then the lesser-included email must be separately produced with the attachments

If a Party produces only inclusive email, the Parties agree to redact any privileged content or message component and produce in full the rest of the conversation in the inclusives.

“Near duplicate” Documents shall be produced rather than removed. The Producing Party need only produce a single copy of a particular ESI.

The hash value for each item shall be reflected in the .DAT load file in the HashValue field specified in Appendix 2. If a Producing Party elects to de-duplicate globally, the Producing Party shall identify custodians who were in possession of a de-duplicated Document in the AllCustodian metadata field specified in Appendix 2. This means that the field “AllCustodian”

will be populated showing all custodians who had a copy of the same Document which is not being produced because of de-duplication.

De-duplication shall not break apart families and shall be performed at a family level. A Document and all other Documents in its attachment range, emails with attachments and files with extracted embedded OLE Documents all constitute family groups, subject to the exclusions identified in I(G). If any member of a family group is produced, all members of that group must be also be produced or else logged and no such member shall be withheld from production as a duplicate. The Producing Party agrees that the presence of a custodian's name contained in the "AllCustodian" field in the metadata for a particular Document is evidence that the custodian possessed that Document in his/her custodial file.

F. Load Files:

The data load file should be in standard Concordance format (.DAT). The .DAT file shall contain a relative path to the corresponding native file.

Concordance Data Load Files:

The data load file should use standard Concordance delimiters:

- Column - ¶ (ASCII 182);
- Quote - ¢ (ASCII 254);
- Newline - ® (ASCII 174).
- The first line of the .DAT file should contain the field names arranged in the same order as the data is arranged in subsequent lines.
- All date fields should be produced in mm/dd/yyyy format, if possible.
Date fields may be combined date/time fields
- All produced attachments should sequentially follow the parent
Document/email.

Sample Concordance .DAT Load File:

bBegBatesb bBegattatchb bCustodianb bNative Pathb bExtracted Textb
 bABC000001b bABC000001b bSampleb b\VOL001\NATIVES\001\ABC000001.xlsb bTEXT
 \001\ABC000001.t
 tb bABC000002b bABC000001b bSampleb bNATIVES\001\ABC000002.gifb bTEXT\001\AB
 C000002.txtb

The image load file must be in standard Option (.OPT) format and must reference each
 TIFF file in the corresponding production, and the total number of TIFF files referenced in the
 load file must match the total number of image files in the production. The total number of
 Documents referenced in a production's data load file should match the total number of
 designated document breaks in the corresponding image load file for that production. In any
 deliverable volume, Documents should be organized in such a way that each folder in the volume
 contains 1000 files (each TIFF page or native file is a single file) as one file per folder.

- Every image in the delivery volume should be cross-referenced in the image load file.
- The imageID key should be named the same as the Bates number of the page.
- Load files should not span across media (e.g., CDs, DVDs, hard drives, etc.), i.e., a separate volume should be created for each piece of media delivered.

Sample Concordance/Opticon Image (.OPT) Load File:

MSC000001, MSC001, \VOL001\IMAGES\001\MSC000001.TIF, Y,,,2

MSC000002, MSC001, \VOL001\IMAGES\001\MSC000002.TIF,,,,

MSC000004, MSC001, \VOL001\IMAGES\001\MSC000004.TIF, Y,,,2

MSC000005, MSC001, \VOL001\IMAGES\001\MSC000005.TIF,,,

OCR/Extracted Text Files:

- OCR or Extracted Text files shall be provided in a separate directory containing Document-level text files. The .DAT file shall contain a relative path to the corresponding text file. OCR or Extracted text itself should not be included in the .DAT file:

PROD001\TEXT\001\ABC00015.txt

APPENDIX 2
ESI METADATA AND CODING FIELDS

The chart below describes the metadata fields to be produced, where reasonably available, in generic, commonly used terms which the Producing Party is to adapt to the specific types of ESI it is producing, to the extent such metadata fields exist associated with the original electronic Documents and are automatically generated as part of the electronic data discovery process. Any ambiguity about a metadata field should be discussed with the Receiving Party prior to processing the subject ESI for production.

Field Name	Field Description
BegBates	First Bates number (production number) of an item
EndBates	Last Bates number (production number) of an item **The EndBates field should be populated for single-page items.
BegAttach/Group ID	First Bates number of family group.
EndAttach	Last Bates number of attachment range (<i>i.e.</i> , Bates number of the last page of the last attachment)
PgCount	Number of pages in the item
Custodian	Name of person or source from whose/which files the item is produced
AllCustodian	Name of the person(s), in addition to the Custodian, from whose files the item would have been produced if it had not been de- duplicated
FileSize	Size (in kilobytes) of the source native file
SourceFilePath ²	The directory structure or path where the original file was stored on the Party's source computer system, ending in the filename. Any container name (such as ZIP or PST containers) is included in the path.

² This field may be excluded if the Producing Party determines it includes information protected by any applicable privilege or immunity.

Field Name	Field Description
HashValue	The MD5 or SHA-1 or IntMsgID hash value of the item.
RecordType	Indicates item type (<i>e.g.</i> , email, edoc, attachment)
FileType	(<i>e.g.</i> , Outlook, Adobe Acrobat, MS Word, etc.)
FileExtension	Indicates file extension of the file <i>e.g.</i> , .docx, .pptx
Parent Date/Sort Date	<p>The date associated with a family's parent record, which is assigned as follows:</p> <p>Emails populated with 1st occurrence: Date Sent or Date Received or Date Last Modified. Email attachments populated from the date above.</p> <p>Outlook Appointments populated with 1st occurrence of Start Date or Date Last Modified.</p> <p>Loose Edocs populated with 1st occurrence of Date Last Modified or Date Created.</p>
DateSent (mm/dd/yyyy)	Date email or calendar item was sent
TimeSent (hh:mmAM/PM)	Time email or calendar item was sent (Date and time fields may be combined)
DateReceived	Date email or calendar item was received
TimeReceived	Time email or calendar item was received (Date and time fields may be combined)
To	The names and/or SMTP email addresses of all recipients that were included on the "To" line of the email or calendar item
From	The name and/or SMTP email address of the sender of the email or calendar item
CC	The names and/or SMTP email addresses of all recipients that were included on the "CC" line of the email or calendar item
BCC	The names and/or SMTP email addresses of all recipients that were included on the "BCC" line of the email or calendar item

Field Name	Field Description
Number of Attachments	Number of attached, embedded or grouped items
DateCreated (mm/dd/yyyy)	Date the item was created
TimeCreated (hh:mm AM/PM)	Time the item was created (Date and time fields may be combined)
ModifiedBy	Person who last modified or saved the item, as populated in the metadata or document properties of the native file
LastModDate (mm/dd/yyyy)	Date the item was last modified
LastModTime (hh:mm AM/PM)	Time the item was last modified
Date Last Printed	Date the Document was last printed.
FileName	The filename of the source native file for an ESI item
Title	Any value populated in the Title field of the source file metadata or item properties
Subject/E-Mail Subject	Any value populated in the Subject field of the source file metadata or document properties (e.g., subject line of email or calendar item)
Author	Creator of the Document; any value populated in the Author field of the source file metadata or document properties
Redacted	User-generated field that will indicate redactions. With the word "REDACTED". Otherwise, blank.
Redaction Reason	The reason for the redaction. If more than one reason, separate by semi-colons.
Confidentiality	User-generated field that will indicate confidentiality. With the word "Confidential" or "Highly Confidential – Attorneys' Eyes Only" applicable. Otherwise, blank.
TextPath	Full relative path to the location of the Document-level text file.

Field Name	Field Description
NativeFileLink	Relative path for Documents provided in Native Format only. **The linked file must be named per the BegBates value and confidentiality designation, where reasonably feasible.